

PCT09

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/830,506

TIME: 16:26:48

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

3 <110> APPLICANT: Genzyme Corporation
5 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
6 AND TREATING POLYCYSTIC KIDNEY DISEASE
9 <130> FILE REFERENCE: 126881206140
C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/830,506
C--> 12 <141> CURRENT FILING DATE: 2001-08-10
14 <150> PRIOR APPLICATION NUMBER: 60/105,731
15 <151> PRIOR FILING DATE: 1998-10-28
17 <150> PRIOR APPLICATION NUMBER: 60/105,876
18 <151> PRIOR FILING DATE: 1998-10-27
20 <150> PRIOR APPLICATION NUMBER: 60/141,175
21 <151> PRIOR FILING DATE: 1999-06-25
23 <160> NUMBER OF SEQ ID NOS: 3
25 <170> SOFTWARE: FastSEQ for Windows Version 3.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 14060
29 <212> TYPE: DNA
30 <213> ORGANISM: Homo sapiens
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (135)...(13040)
36 <221> NAME/KEY: misc_feature
37 <222> LOCATION: (2621)...(2710)
38 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
39 transmembrane domain
41 <221> NAME/KEY: misc_feature
42 <222> LOCATION: (2734)...(3094)
43 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
44 transmembrane domain
46 <221> NAME/KEY: misc_feature
47 <222> LOCATION: (3166)...(3300)
48 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
49 transmembrane domain
51 <221> NAME/KEY: misc_feature
52 <222> LOCATION: (3364)...(3578)
53 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
54 transmembrane domain
56 <221> NAME/KEY: misc_feature
57 <222> LOCATION: (3623)...(3688)
58 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
59 transmembrane domain
61 <221> NAME/KEY: misc_feature
62 <222> LOCATION: (3710)...(3914)
63 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
64 transmembrane domain
66 <221> NAME/KEY: misc_feature
67 <222> LOCATION: (3931)...(4046)

ENTERED

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/830,506

TIME: 16:26:48

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

```

68 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
69     transmembrane domain
71 <221> NAME/KEY: misc_feature
72 <222> LOCATION: (2166)...(2599)
73 <223> OTHER INFORMATION: epitope outside the loop region but within the
74     polycystin transmembrane domain
76 <221> NAME/KEY: misc_feature
77 <222> LOCATION: (4097)...(4302)
78 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
79     transmembrane domain
81 <221> NAME/KEY: misc_feature
82 <222> LOCATION: (4148)...(4219)
83 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
84     transmembrane domain
86 <221> NAME/KEY: misc_feature
87 <222> LOCATION: (4220)...(4302)
88 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
89     transmembrane domain
91 <221> NAME/KEY: misc_feature
92 <222> LOCATION: (27)...(360)
93 <223> OTHER INFORMATION: epitope in the loop region of the polycystin
94     transmembrane domain
96 <221> NAME/KEY: misc_feature
97 <222> LOCATION: (843)...(1200)
98 <223> OTHER INFORMATION: Ig-like domain of polycystin
100 <221> NAME/KEY: misc_feature
101 <222> LOCATION: (1205)...(1625)
102 <223> OTHER INFORMATION: Ig-like domain of polycystin
104 <221> NAME/KEY: misc_feature
105 <222> LOCATION: (1626)...(2136)
106 <223> OTHER INFORMATION: Ig-like domain of polycystin
108 <221> NAME/KEY: misc_feature
109 <222> LOCATION: (2166)...(2599)
110 <223> OTHER INFORMATION: isolated polypeptide
112 <400> SEQUENCE: 1
113 gctcagcagc aggtcgcggc cgcagcccca tccagcccgc gcccgccatg ccgtccgcgg      60
114 gccccgcctg agctgcggtc tccgcgcgcg ggcgggctg gggacggcgg ggccatgcgc      120
115 gcgctgccct aacg atg ccg ccc gcc gcg ccc gcc cgc ctg gcg ctg gcc      170
116             Met Pro Pro Ala Ala Pro Ala Arg Leu Ala Leu Ala
117             1             5             10
119 ctg ggc ctg ggc ctg tgg ctc ggg gcg ctg gcg ggg ggc ccc ggg cgc      218
120 Leu Gly Leu Gly Leu Trp Leu Gly Ala Leu Ala Gly Gly Pro Gly Arg
121             15             20             25
123 ggc tgc ggg ccc tgc gag ccc ccc tgc ctc tgc ggc cca gcg ccc ggc      266
124 Gly Cys Gly Pro Cys Glu Pro Pro Cys Leu Cys Gly Pro Ala Pro Gly
125             30             35             40
127 gcc gcc tgc cgc gtc aac tgc tcg ggc cgc ggg ctg cgg acg ctc ggt      314
128 Ala Ala Cys Arg Val Asn Cys Ser Gly Arg Gly Leu Arg Thr Leu Gly
129             45             50             55             60

```

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/830,506

TIME: 16:26:48

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

131	ccc gcg ctg cgc atc ccc gcg gac gcc aca gcg cta gac gtc tcc cac	362
132	Pro Ala Leu Arg Ile Pro Ala Asp Ala Thr Ala Leu Asp Val Ser His	
133	65 70 75	
135	aac ctg ctc cgg gcg ctg gac gtt ggg ctc ctg gcg aac ctc tcg gcg	410
136	Asn Leu Leu Arg Ala Leu Asp Val Gly Leu Leu Ala Asn Leu Ser Ala	
137	80 85 90	
139	ctg gca gag ctg gat ata agc aac aac aag att tct acg tta gaa gaa	458
140	Leu Ala Glu Leu Asp Ile Ser Asn Asn Lys Ile Ser Thr Leu Glu Glu	
141	95 100 105	
143	gga ata ttt gct aat tta ttt aat tta agt gaa ata aac ctg agt ggg	506
144	Gly Ile Phe Ala Asn Leu Phe Asn Leu Ser Glu Ile Asn Leu Ser Gly	
145	110 115 120	
147	aac ccg ttt gag tgt gac tgt ggc ctg gcg tgg ctg ccg cga tgg gcg	554
148	Asn Pro Phe Glu Cys Asp Cys Gly Leu Ala Trp Leu Pro Arg Trp Ala	
149	125 130 135 140	
151	gag gag cag cag gtg cgg gtg gtg cag ccc gag gca gcc acg tgt gct	602
152	Glu Glu Gln Gln Val Arg Val Val Gln Pro Glu Ala Ala Thr Cys Ala	
153	145 150 155	
155	ggg cct ggc tcc ctg gct ggc cag cct ctg ctt ggc atc ccc ttg ctg	650
156	Gly Pro Gly Ser Leu Ala Gly Gln Pro Leu Leu Gly Ile Pro Leu Leu	
157	160 165 170	
159	gac agt ggc tgt ggt gag gag tat gtc gcc tgc ctc cct gac aac agc	698
160	Asp Ser Gly Cys Gly Glu Glu Tyr Val Ala Cys Leu Pro Asp Asn Ser	
161	175 180 185	
163	tca ggc acc gtg gca gca gtg tcc ttt tca gct gcc cac gaa ggc ctg	746
164	Ser Gly Thr Val Ala Ala Val Ser Phe Ser Ala Ala His Glu Gly Leu	
165	190 195 200	
167	ctt cag cca gag gcc tgc agc gcc ttc tgc ttc tcc acc ggc cag ggc	794
168	Leu Gln Pro Glu Ala Cys Ser Ala Phe Cys Phe Ser Thr Gly Gln Gly	
169	205 210 215 220	
171	ctc gca gcc ctc tcg gag cag ggc tgg tgc ctg tgt ggg gcg gcc cag	842
172	Leu Ala Ala Leu Ser Glu Gln Gly Trp Cys Leu Cys Gly Ala Ala Gln	
173	225 230 235	
175	ccc tcc agt gcc tcc ttt gcc tgc ctg tcc ctc tgc tcc ggc ccc ccg	890
176	Pro Ser Ser Ala Ser Phe Ala Cys Leu Ser Leu Cys Ser Gly Pro Pro	
177	240 245 250	
179	cca cct cct gcc ccc acc tgt agg ggc ccc acc ctc ctc cag cac gtc	938
180	Pro Pro Pro Ala Pro Thr Cys Arg Gly Pro Thr Leu Leu Gln His Val	
181	255 260 265	
183	ttc cct gcc tcc cca ggg gcc acc ctg gtg ggg ccc cac gga cct ctg	986
184	Phe Pro Ala Ser Pro Gly Ala Thr Leu Val Gly Pro His Gly Pro Leu	
185	270 275 280	
187	gcc tct ggc cag cta gca gcc ttc cac atc gct gcc ccg ctc cct gtc	1034
188	Ala Ser Gly Gln Leu Ala Ala Phe His Ile Ala Ala Pro Leu Pro Val	
189	285 290 295 300	
191	act gcc aca cgc tgg gac ttc gga gac ggc tcc gcc gag gtg gat gcc	1082
192	Thr Ala Thr Arg Trp Asp Phe Gly Asp Gly Ser Ala Glu Val Asp Ala	
193	305 310 315	
195	gct ggg ccg gct gcc tcg cat cgc tat gtg ctg cct ggg cgc tat cac	1130

RAW SEQUENCE LISTING

DATE: 09/24/2001

PATENT APPLICATION: US/09/830,506

TIME: 16:26:48

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

196	Ala	Gly	Pro	Ala	Ala	Ser	His	Arg	Tyr	Val	Leu	Pro	Gly	Arg	Tyr	His	
197				320					325					330			
199	gtg	acg	gcc	gtg	ctg	gcc	ctg	ggg	gcc	ggc	tca	gcc	ctg	ctg	ggg	aca	1178
200	Val	Thr	Ala	Val	Leu	Ala	Leu	Gly	Ala	Gly	Ser	Ala	Leu	Leu	Gly	Thr	
201			335					340					345				
203	gac	gtg	cag	gtg	gaa	gcg	gca	cct	gcc	gcc	ctg	gag	ctc	gtg	tgc	ccg	1226
204	Asp	Val	Gln	Val	Glu	Ala	Ala	Pro	Ala	Ala	Leu	Glu	Leu	Val	Cys	Pro	
205		350					355					360					
207	tcc	tcg	gtg	cag	agt	gac	gag	agc	ctt	gac	ctc	agc	atc	cag	aac	cgc	1274
208	Ser	Ser	Val	Gln	Ser	Asp	Glu	Ser	Leu	Asp	Leu	Ser	Ile	Gln	Asn	Arg	
209	365					370				375					380		
211	ggg	ggg	tca	ggc	ctg	gag	gcc	gcc	tac	agc	atc	gtg	gcc	ctg	ggc	gag	1322
212	Gly	Gly	Ser	Gly	Leu	Glu	Ala	Ala	Tyr	Ser	Ile	Val	Ala	Leu	Gly	Glu	
213				385					390				395				
215	gag	ccg	gcc	cga	gcg	gtg	cac	ccg	ctc	tgc	ccc	tcg	gac	acg	gag	atc	1370
216	Glu	Pro	Ala	Arg	Ala	Val	His	Pro	Leu	Cys	Pro	Ser	Asp	Thr	Glu	Ile	
217			400					405					410				
219	ttc	cct	ggc	aac	ggg	cac	tgc	tac	cgc	ctg	gtg	gtg	gag	aag	gcg	gcc	1418
220	Phe	Pro	Gly	Asn	Gly	His	Cys	Tyr	Arg	Leu	Val	Val	Glu	Lys	Ala	Ala	
221		415					420					425					
223	tgg	ctg	cag	gcg	cag	gag	cag	tgt	cag	gcc	tgg	gcc	ggg	gcc	gcc	ctg	1466
224	Trp	Leu	Gln	Ala	Gln	Glu	Gln	Cys	Gln	Ala	Trp	Ala	Gly	Ala	Ala	Leu	
225		430				435					440						
227	gca	atg	gtg	gac	agt	ccc	gcc	gtg	cag	cgc	ttc	ctg	gtc	tcc	cgg	gtc	1514
228	Ala	Met	Val	Asp	Ser	Pro	Ala	Val	Gln	Arg	Phe	Leu	Val	Ser	Arg	Val	
229	445				450				455				460				
231	acc	agg	agc	cta	gac	gtg	tgg	atc	ggc	ttc	tcg	act	gtg	cag	ggg	gtg	1562
232	Thr	Arg	Ser	Leu	Asp	Val	Trp	Ile	Gly	Phe	Ser	Thr	Val	Gln	Gly	Val	
233			465					470				475					
235	gag	gtg	ggc	cca	gcg	ccg	cag	ggc	gag	gcc	ttc	agc	ctg	gag	agc	tgc	1610
236	Glu	Val	Gly	Pro	Ala	Pro	Gln	Gly	Glu	Ala	Phe	Ser	Leu	Glu	Ser	Cys	
237		480					485					490					
239	cag	aac	tgg	ctg	ccc	ggg	gag	cca	cac	cca	gcc	aca	gcc	gag	cac	tgc	1658
240	Gln	Asn	Trp	Leu	Pro	Gly	Glu	Pro	His	Pro	Ala	Thr	Ala	Glu	His	Cys	
241		495				500			505				510				
243	gtc	cgg	ctc	ggg	ccc	acc	ggg	tgg	tgt	aac	acc	gac	ctg	tgc	tca	gcg	1706
244	Val	Arg	Leu	Gly	Pro	Thr	Gly	Trp	Cys	Asn	Thr	Asp	Leu	Cys	Ser	Ala	
245		510				515			520								
247	ccg	cac	agc	tac	gtc	tgc	gag	ctg	cag	ccc	gga	ggc	cca	gtg	cag	gat	1754
248	Pro	His	Ser	Tyr	Val	Cys	Glu	Leu	Gln	Pro	Gly	Gly	Pro	Val	Gln	Asp	
249	525				530				535				540				
251	gcc	gag	aac	ctc	ctc	gtg	gga	gcg	ccc	agt	ggg	gac	ctg	cag	gga	ccc	1802
252	Ala	Glu	Asn	Leu	Leu	Val	Gly	Ala	Pro	Ser	Gly	Asp	Leu	Gln	Gly	Pro	
253			545				550					555					
255	ctg	acg	cct	ctg	gca	cag	cag	gac	ggc	ctc	tca	gcc	ccg	cac	gag	ccc	1850
256	Leu	Thr	Pro	Leu	Ala	Gln	Gln	Asp	Gly	Leu	Ser	Ala	Pro	His	Glu	Pro	
257		560					565					570					
259	gtg	gag	gtc	atg	gta	ttc	ccg	ggc	ctg	cgt	ctg	agc	cgt	gaa	gcc	ttc	1898
260	Val	Glu	Val	Met	Val	Phe	Pro	Gly	Leu	Arg	Leu	Ser	Arg	Glu	Ala	Phe	

DATE: 09/24/2001

TIME: 16:26:48

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

261	575						580						585						
263	ctc	acc	acg	gcc	gaa	ttt	ggg	acc	cag	gag	ctc	cgg	cgg	ccc	gcc	cag	1946		
264	Leu	Thr	Thr	Ala	Glu	Phe	Gly	Thr	Gln	Glu	Leu	Arg	Arg	Pro	Ala	Gln			
265	590						595						600						
267	ctg	cgg	ctg	cag	gtg	tac	cgg	ctc	ctc	agc	aca	gca	ggg	acc	ccg	gag	1994		
268	Leu	Arg	Leu	Gln	Val	Tyr	Arg	Leu	Leu	Ser	Thr	Ala	Gly	Thr	Pro	Glu			
269	605						610						615						620
271	aac	ggc	agc	gag	cct	gag	agc	agg	tcc	ccg	gac	aac	agg	acc	cag	ctg	2042		
272	Asn	Gly	Ser	Glu	Pro	Glu	Ser	Arg	Ser	Pro	Asp	Asn	Arg	Thr	Gln	Leu			
273	625						630						635						
275	gcc	ccc	gcg	tgc	atg	cca	ggg	gga	cgc	tgg	tgc	cct	gga	gcc	aac	atc	2090		
276	Ala	Pro	Ala	Cys	Met	Pro	Gly	Gly	Arg	Trp	Cys	Pro	Gly	Ala	Asn	Ile			
277	640						645						650						
279	tgc	ttg	ccg	ctg	gac	gcc	tct	tgc	cac	ccc	cag	gcc	tgc	gcc	aat	ggc	2138		
280	Cys	Leu	Pro	Leu	Asp	Ala	Ser	Cys	His	Pro	Gln	Ala	Cys	Ala	Asn	Gly			
281	655						660						665						
283	tgc	acg	tca	ggg	cca	ggg	cta	ccc	ggg	gcc	ccc	tat	gcg	cta	tgg	aga	2186		
284	Cys	Thr	Ser	Gly	Pro	Gly	Leu	Pro	Gly	Ala	Pro	Tyr	Ala	Leu	Trp	Arg			
285	670						675						680						
287	gag	ttc	ctc	ttc	tcc	gtt	gcc	gcg	ggg	ccc	ccc	gcg	cag	tac	tcg	gtc	2234		
288	Glu	Phe	Leu	Phe	Ser	Val	Ala	Ala	Gly	Pro	Pro	Ala	Gln	Tyr	Ser	Val			
289	685						690						695						700
291	acc	ctc	cac	ggc	cag	gat	gtc	ctc	atg	ctc	cct	ggt	gac	ctc	gtt	ggc	2282		
292	Thr	Leu	His	Gly	Gln	Asp	Val	Leu	Met	Leu	Pro	Gly	Asp	Leu	Val	Gly			
293	705						710						715						
295	ttg	cag	cac	gac	gct	ggc	cct	ggc	gcc	ctc	ctg	cac	tgc	tcg	ccg	gct	2330		
296	Leu	Gln	His	Asp	Ala	Gly	Pro	Gly	Ala	Leu	Leu	His	Cys	Ser	Pro	Ala			
297	720						725						730						
299	ccc	ggc	cac	cct	ggt	ccc	cag	gcc	ccg	tac	ctc	tcc	gcc	aac	gcc	tcg	2378		
300	Pro	Gly	His	Pro	Gly	Pro	Gln	Ala	Pro	Tyr	Leu	Ser	Ala	Asn	Ala	Ser			
301	735						740						745						
303	tca	tgg	ctg	ccc	cac	ttg	cca	gcc	cag	ctg	gag	ggc	act	tgg	gcc	tgc	2426		
304	Ser	Trp	Leu	Pro	His	Leu	Pro	Ala	Gln	Leu	Glu	Gly	Thr	Trp	Ala	Cys			
305	750						755						760						
307	cct	gcc	tgt	gcc	ctg	cgg	ctg	ctt	gca	gcc	acg	gaa	cag	ctc	acc	gtg	2474		
308	Pro	Ala	Cys	Ala	Leu	Arg	Leu	Leu	Ala	Ala	Thr	Glu	Gln	Leu	Thr	Val			
309	765						770						775						780
311	ctg	ctg	ggc	ttg	agg	ccc	aac	cct	gga	ctg	cgg	atg	cct	ggg	cgc	tat	2522		
312	Leu	Leu	Gly	Leu	Arg	Pro	Asn	Pro	Gly	Leu	Arg	Met	Pro	Gly	Arg	Tyr			
313	785						790						795						
315	gag	gtc	cgg	gca	gag	gtg	ggc	aat	ggc	gtg	tcc	agg	cac	aac	ctc	tcc	2570		
316	Glu	Val	Arg	Ala	Glu	Val	Gly	Asn	Gly	Val	Ser	Arg	His	Asn	Leu	Ser			
317																			

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/830,506

DATE: 09/24/2001

TIME: 16:26:49

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\09242001\I830506.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date